John von Neumann Institute for Computing

First-Principles Molecular Dynamics Simulations of Liquids and Solvent Effects

J. Thar, B. Kirchner

published in

NIC Symposium 2008,
G. Münster, D. Wolf, M. Kremer (Editors),
John von Neumann Institute for Computing, Jülich,
NIC Series, Vol. **39**, ISBN 978-3-9810843-5-1, pp. 87-94, 2008.

http://www.fz-juelich.de/nic-series/volume39

First-Principles Molecular Dynamics Simulations of Liquids and Solvent Effects

Jens Thar and Barbara Kirchner

Wilhelm-Ostwald Institut für Physikalische und Theoretische Chemie
Universität Leipzig, Linnestr. 2, D-04103 Leipzig
E-mail: bkirchner@uni-leipzig.de

We present several *first-principles* molecular dynamics simulations studies of liquids and solvent effects. The first project deals with dipole moments of one alanine molecule dissolved in water molecules. The average dipole moment of the water molecules depends on their vicinity to the distinct functional groups of alanine. The results of the *first-principles* molecular dynamics simulation is used to verify a cluster ansatz. In this ansatz, snapshots of alanine surrounded by different shells of water molecules are cut out of a trajectory. Sufficient convergence of the dipole moment of alanine and its first water shell can be found for two solvent shells. In the second project, we examine the basis set superposition error along a free-energy trajectory of the water dissociation reaction. Snapshots are again cut out of a trajectory and their basis set superposition error is calculated by static quantum mechanics. We observe an increase of the basis set superposition error with the decline of the distance of the two molecules. In the third project which is still in progress, we investigate an example of the third generation ionic liquid, ethyl-methyl-imidazoliumthiocyanate ($[C_2C_1im]SCN$).

1 Introduction

1.1 Alanine in Water[1]

The solvation behaviour of amino acids in water is one of the key questions to be addressed in biology, due to its drastic influence on their chemical reactivity. The most straightforward way to adequately describe such a flexible system are *first-principles* molecular dynamics simulations. In this project the electronic structure is accurately calculated using density functional theory, while the configurational space is explored by an efficient and temperature dependent sampling technique.

Another possibility to derive properties of molecules dissolved in liquids is the so called cluster ansatz where a trajectory either obtained by classical or first-principles molecular dynamics simulations samples the configurations in the solvents. This trajectory is subsequently harvested for small clusters. The main concern with this approach is the number of snapshots and the number of solvent molecules needed to achieve convergence with respect to the experiment.

1.2 BSSE Error along the Free Energy Trajectory[2]

Atom-centreed basis sets (e.g. the Gaussian Basis Set (GBS)) that are used in static quantum calculations introduce the so-called basis set superposition error (BSSE), when comparing interaction energies (E_I) at different geometries. The effect of this error on the static geometries is under control, because it remains nearly constant close to the minimum structure. In simulations, the average structure is determined by sampling a broader

region around the local minima and by weighting configurations according to their energy. Therefore, the BSSE is not necessarily constant, and might have a direct influence on the weighting and hence the average structure.

1.3 Simulation of Third Generation Ionic Liquids

The wide applicability of ionic liquids as environmentally friendly solvents has led to a steep increase of publications on this subject. Ionic liquids are fused salts of organic cations and most often inorganic anions.[3,4] They are good solvents for a wide range of inorganic and organic reactions. The reason for this is attributed to their microscopic properties. These properties are mainly their high polarity and their non-coordinating behaviour.

2 Scope

2.1 Alanine in Water

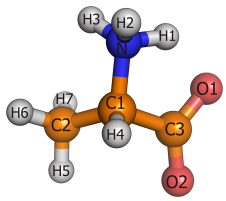

Figure 1. Structure of alanine with atom labels as used in this chapter.

In this project, we investigate alanine as the smallest biologically active α-amino acid. The zwitterionic form of the alanine shows three sites with distinct polarity: The basic COO^- group, the acidic NH_3^+ group, and the aliphatic CH_3 group. The main focus is the development of the dipole moment of a zwitterionic alanine in water as well as of different water shells along a trajectory. These local dipole moments are estimated by *first-principles* simulation in combination with maximally localized Wannier centres (MLWC). Additionally, we want to test to what extent the cluster ansatz not relying on periodic boundary conditions and with fewer solvent molecules produces comparable results. Therefore, we cut out snapshots with different numbers of solvent molecules along the simulated trajectory and calculate local dipole moments by means of MLWCs. This allows us to verify the cluster ansatz against the results from our previous *first principles* simulation.

2.2 BSSE Error along the Free Energy Trajectory

We investigate the reaction

$$\mathrm{H_2O^\star \cdots H^\star\text{–}O\text{–}H \rightarrow H_3O^+ + OH^-}$$

by means of thermodynamical integration. Despite the fact that this reaction does not take place in the gas phase, it serves us to obtain many chemically different situations. During its course, each of these situations would be associated with a unique BSSE when a Gaussian type orbital basis set would be applied. The trajectory is therefore harvested for these different configurations, and their BSSE is determined by means of static quantum mechanics.

2.3 Simulation of Third Generation Ionic Liquids

Here, we investigate a third generation ionic liquid, ethyl-methylimidazoliumthiocyanate ($[C_2C_1im]SCN$).

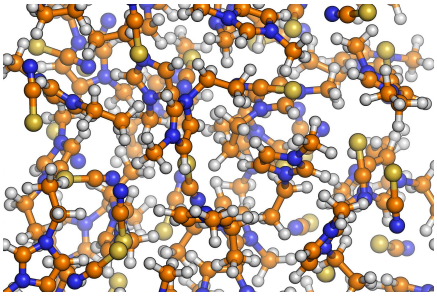

Figure 2. Structure of ethyl-methylimidazoliumthiocyanate.

A *first-principles* molecular dynamics simulation can determine the main building blocks of the liquid and their life time as well as the diffusion constant of the dominating species.

3 Methodology

We simulate one alanine as well as the free energy trajectory with the CPMD code, which is based on the Car–Parrinello method [6]. The electronic structure is described in the framework of density functional theory, utilising the gradient corrected exchange correlation functional BP86. The Kohn–Sham orbitals are expanded in a plane wave (PW) basis with a kinetic energy cutoff of 70 Rydberg, while core electrons are approximated by norm conserving pseudo potentials of the Troullier–Martins type[5]. Local dipole moments for CPMD calculations are derived based on maximally localised Wannier centres (MLWC)s.

For alanine, the system consists of 60 water molecules surrounding the alanine in a periodic box. The equilibration simulation was thermostated at 320 K using Nosé–Hoover thermostats[7–9], while the production run was performed in the NVE ensemble. Average computer time for a Car–Parrinello molecular dynamics step on 16 CPUs of the JUMP Cluster was 7.4 seconds. Additional averaged time needed for calculation of the maximally localized Wannier centres (MLWC)s was 14.8 seconds.

Concerning the BSSE, this error is estimated by the counterpoise correction (CP) of the dissociation reaction given in Eq. (2.2). Each monomer m is calculated in the basis set

centred at its atoms (E_m^{mon}) and in the complete basis set (E_m^{all}). The CP of the reaction is defined by the difference $(E_1^{mon} + E_2^{mon}) - (E_1^{all} + E_2^{all})$.

For $[C_2C_1im]SCN$, molecular dynamics simulation using the GPW[10] ansatz in the framework of the CP2K/Quickstep[11] environment has been performed. Temperature was set to 398.15 K and controlled by a Nosé–Hoover thermosthat. We use GTH type pseudopotentials[12,13] to describe the core electrons of all atoms. A DZVP basis set is used for the representation of the Kohn–Sham orbitals, and upon switching to the plane wave representation, a cutoff of 300 Ryd is used. We use the gradient corrected functional PBE throughout the simulation.

4 Results: Alanin

4.1 Dipole Moments from Molecular Dynamics: Water

We investigate the dipole moment of water molecules around each of the three distinct functional groups NH_3^+, CH_3 and COO^-. The probability distributions of the dipole moment from different shells inside each of these cones is depicted in Fig. 3. The first shell around the oxygen atom contains all water molecules up to a distance of 330 pm, the second shell contains all water molecules from 330 pm up to a distance of 540 pm, and further solvent molecules are collected in the third shell. For nitrogen, these two cutoffs are set to 360 pm and 590 pm, and for the carbon atoms, the first shell consists of water molecules up to a distance of 510 pm, and all further solvent molecules are collected in the second shell.

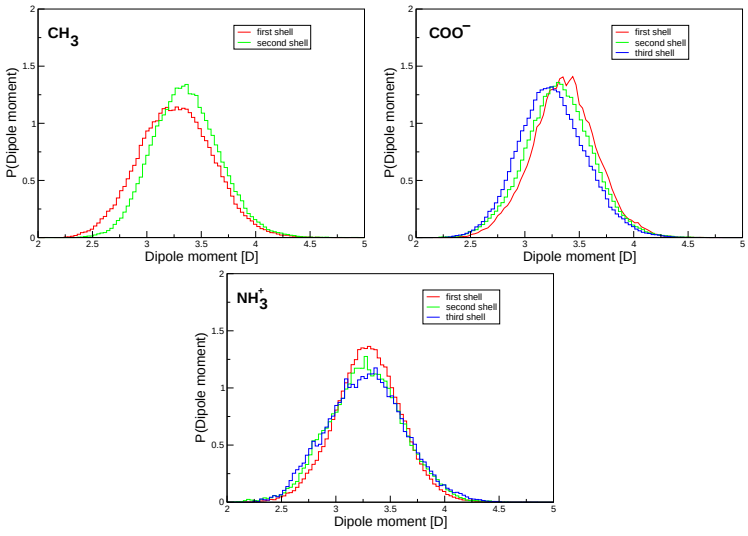

Figure 3. Probability distribution function of the dipole moments of water molecules closest to atoms of the COO^-, the CH_3 and the $NH3^+$ group of the alanine. See text for details.

As expected, the average dipole moment of the solvent molecules in the vincity of the

COO^- decays from 3.35 D (first shell) over 3.31 D (second shell) to 3.25 D (third shell) with rising distance from the solute. In contrast, the average dipole moments of the solvent molecules around the CH_3 and the NH_3 group of alanine rise from 3.27 D (3.28 D) to 3.36 D (3.0 D) when entering the second shell.

4.2 Dipole Moments from Static Calculations: The Cluster-Ansatz

In this project we want to explore whether the use of periodic boundary conditions is necessary for the computation of electronic properties of the solute (i.e. the alanine) or if these properties are sufficiently converged already with two or even less layers of solvent molecules added. i.e., we want to investigate for which cluster size the cluster method is valid.

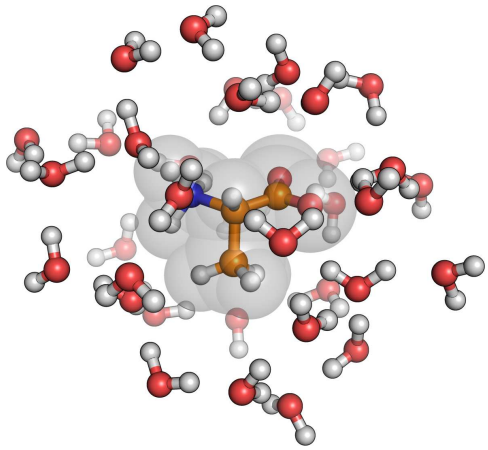

Figure 4. Snapshot of one set solvent molecules.

Therefore, we selected 4 distinct sets of snapshots: The first two sets only account for the first (and second) solvent shell of the two polar groups. The second two sets account for the first (and second) solvent shell of all selected subgroups. A short assay revealed that already 50 snapshots per set are sufficient to converge the average dipole moments.

Method	**set0**	**set1**	**set2**	**set3**	**set4**	**full trajectory**
PW/MLWC	10.99	14.10	15.08	14.59	15.56	15.93

Table 1. Average local dipole moments in D of alanine within different sets of solvent molecules.

The influence of different solvent shells on the local dipole moment of alanine is summarized in table 1. The dipole moment converges nicely with increasing number of electrons. As a result, **set4** is already a reasonable good approximation to the periodic boundary conditions, yielding about 98% of the reference dipole moment. A similar analysis

has been performed for the dipole moment of the water molecules around the first shell of the alanine. We obtain a dipole moment of 3.15 D, about 95% of the reference simulation's value, for the innermost water molecules in **set4**. Thus, the cluster ansatz provides a reliable tool to assess properties of systems with a large number of degrees of freedom.

5 Results: BSSE Error along the Free Energy Trajectory

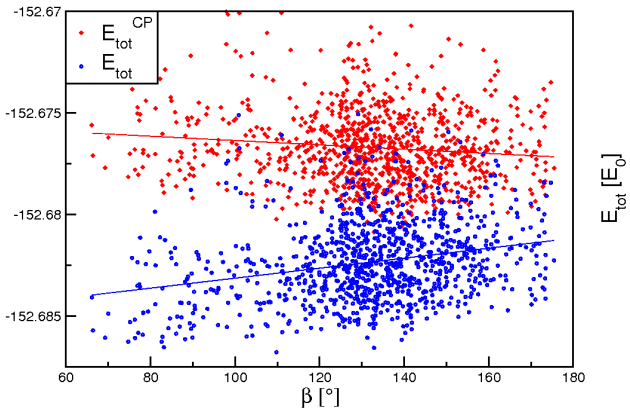

Figure 5. Total energies along the constraint trajectory r1=180 pm plotted against the angle β E_{tot}: uncorrected total SVP energy; E_{tot}^{CP}: counterpoise corrected total SVP energy.

In Fig. 5 we plotted the total energies as well as the total energies with subtracted counterpoise corrections against the angle β for the part of the free energy simulation where the distance between the two products is restrained to 180 pm. β denotes the angle between the vector pointing from the oxygen atom of the first to the oxygen atom of the second water molecule and the vector pointing from the oxygen of the second water molecule to its centre of mass. The corrected values (red) are higher in the graph than the uncorrected values (blue). The straight lines in Fig. 5 show linear regressions in order to clarify the tendencies. It is apparent from the Figure 5 that configurations with small distances between atoms of different molecules (represented here by small angles β) would be overpopulated since their energy (blue) is more favourable than if correct energies (red) would be calculated, i.e., the BSSE is not a simple shift of the potential energy surface.

6 Results: Structure of Bulk ([C_2C_1im]SCN)

In Fig. 6 the radial pair distribution function g(r) of different atoms of the cation ethyl-methyl-imidazolium is given. Namely, the distances between the nitrogen and sulfur of the thiocyanate and the imidazolium ring's three protons of the cation are plotted. The exact definition of each atom is displayed in Fig. 2. We see that the first peak at about 220 pm, representing the smallest distance between the two ions, is part of the $H_2 - N$ g-function.

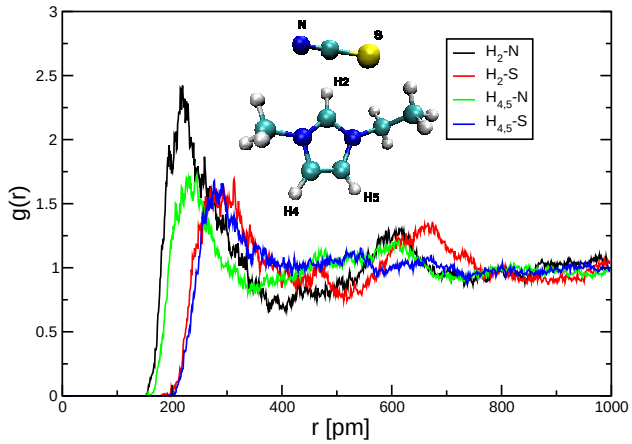

Figure 6. Selected g-function of $[C_2C_1im]SCN$.

The second peak at about 230 pm is provided by the $H_{4,5} - N$ radial pair distribution function. Therefore, we conclude that the preferred coordination site of the cation is the H_2 and of the thiocyanate it is the nitrogen atom. The first "shell" of sulfur atoms around the cation is found at distances of about 280 pm ($H_{4,5} - S$) and 290 pm ($H_2 - S$). It is interesting that this order is opposed to the order found for the nitrogen atoms binding to H_2 or $H_{4,5}$. We must note that according to the average over the autocorrelation function for the orientation of all SCN^- anions' S-N vectors, the simulation is still not converged, and the data presented is still preliminary.

7 Conclusion

We have presented the results of various projects in the field of *first principles* molecular dynamics simulations. Concerning alanine in water, we found that the large dipole moment of alanine induces a considerable increase of the dipole moments of the surrounding water molecules. Also, the different functional groups of alanine exert different but clearly observable influence on their surrounding water molecules. Our test with the cluster ansatz showed that at least 40 water molecules have to be included to converge properties like the dipole moment. Our second project, the impact of the BSSE on simulations using Gaussian basis sets, showed that a sizable BSSE should result in an overestimation of structures which are relatively close together. The third project still runs on JUMP and additional data might be available at the end of this accounting period.

8 Performance on NIC

As a measurement of the performance of the NIC cluster for CP2K, we plotted the average time step of one of our simulations of $[C_2C_1im]SCN$ multiplied with the number of used CPUs against this number, see Fig. (7). The graph shows that CP2K scales favorably for up to 64 CPUs. Using one 16 CPUs is the most time efficient case.

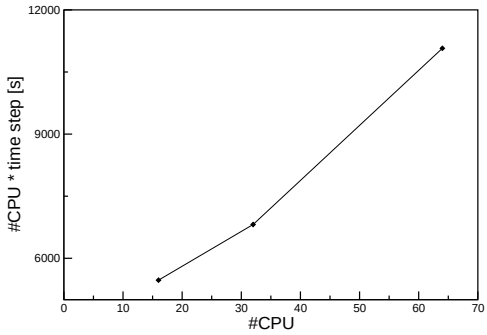

Figure 7. Number of CPUs times the average time step for a CP2K run ploted against the number of CPUs.

Acknowledgments

The authors gratefully acknowledge the financial support of the DFG priority program SPP 1191 “Ionic Liquids” and the ERA Chemistry program, that allows fruitful collaboration under the project “A Modular Approach to Multi-responsive Surfactant/Peptide (SP) and Surfactant/Peptide/Nanoparticle (SPN) Hybrid Materials”. We furthermore like to thank the collaborative research center SFB 624 “Templates” at the University of Bonn for financial support. This work was supported with computational time provided by the NIC supercomputers in Jülich. J.T. thanks funding by a “Chemiefonds-Stipendium” of the Fonds der Chemischen Industrie.

References

1. Thar, J.; Zahn, S.; Kirchner, B. *J. Chem. Phys. B, in print*.
2. Thar, J.; Hovorka, R.; Kirchner, B. *J. Chem. Theor. Comp.* **2007,** *3,* 1510.
3. T. Welton. *Chem. Rev.* **99** (1999) 2071.
4. T. Welton. *Coord. Chem. Rev.* **248** (2004) 2459.
5. Troullier, N.; Martins, J. L. *Phys. Rev. B* **1991,** *43,* 1993.
6. Car, R.; Parrinello, M. *Phys. Rev. Lett.* **1985,** *55,* 2471.
7. Nosé, S. *J. Chem. Phys.* **1984,** *81,* 511.
8. Martyna, G. J.; Klein, M. L.; Tuckerman, M. E. *J. Chem. Phys.* **1992,** *97,* 2635.
9. Hoover, W. G. *Phys. Rev. A* **1985,** *31,* 1695 - 1697.
10. Lippert, G.; Hutter, J.; Parrinello, M. *Theor. Chem. Acc.* **103** (1999) 124.
11. VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J. *Comp. Phys. Comm.* **2005,** *167,* 103.
12. Hartwigsen, C.; Goedecker, S.; Hutter, J. *Phys. Rev. B* **58** (1998) 3641.
13. Goedecker, S.; Teter, M.; Hutter, J. *Phys. Rev. B* **54** (1996) 1703.